MACHINE LEARNING-DRIVEN MAINTENANCE COST OPTIMIZATION: INSIGHTS FROM A LOCAL INDUSTRIAL COMPRESSOR CASE STUDY

[1]Igbokwe Nkemakonam Chidiebube, [1]Fredrick Nnaemeka Okeagu, [1]Nwamekwe Charles Onyeka, [2]Justin Blessing Onwuliri, [1]Ono Chukwuma Godfrey

[1]Industrial/Production Engineering Department, Nnamdi Azikiwe University, P.M.B. 5025 Awka, Anambra State - Nigeria.

[2]Building Department, Faculty of Environmental Sciences, Nnamdi Azikiwe University, P.M.B. 5025 Awka, Anambra State, Nigeria.

*Corresponding author, email: cg.ono@unizik.edu.ng

doi: 10.17977/um068.v4.i11.2024.2

Keywords
Predictive Maintenance
Machine Learning
Industrial Compressors
Maintenance Cost Optimization
Industry 4.0

Abstract
This research explores the application of machine learning (ML) techniques in predictive maintenance for industrial 5-stage compressors, focusing on cost-effectiveness, model performance, and practical implementation. The study evaluates four ML models Random Forest, Decision Trees, Logistic Regression, and Gradient Boosting using real-world data. The Random Forest model demonstrates superior performance with a 94% accuracy, followed closely by Decision Trees. Logistic Regression, while computationally efficient, underperforms in predictive accuracy. Cross-validation and hyperparameter optimization further confirm the Random Forest model's strong generalization capabilities. Cost analysis reveals significant financial benefits from implementing ML-based predictive maintenance, with reductions in downtime and optimized maintenance schedules outweighing the initial investment. The study also highlights the integration of ML into the broader framework of Industry 4.0, emphasizing its potential to enhance equipment reliability, reduce operational costs, and foster intelligent, data-driven maintenance strategies. The research provides a comprehensive framework for industries transitioning from traditional maintenance to ML-driven approaches, contributing to improved operational efficiency and sustainability in industrial environments.

1. Introduction

The increasing demand for efficient maintenance strategies in industrial settings is driven by the need to reduce operational costs while ensuring high reliability and efficiency. Traditional maintenance approaches, such as reactive and time-based strategies, are becoming less sustainable due to the high costs associated with unplanned downtime and the inefficiencies inherent in fixed schedules (Monye, 2023; Okeagu, Nwamekwe, & Nnamani, 2024). In this context, machine learning (ML) emerges as a transformative solution, enabling the development of predictive maintenance strategies that leverage historical data to forecast equipment failures, thus optimizing maintenance interventions (Monye, 2023; Okeagu et al., 2024; Chidiebube, Onyeka, Sunday, & Chiedu, 2025).

This research specifically focuses on applying ML techniques to enhance the maintenance of industrial 5-stage compressors. By utilizing advanced data analytics, industries can transition from conventional maintenance paradigms to more proactive, data-driven approaches that significantly reduce costs and improve operational efficiency (Monye, 2023; Okeagu et al., 2024; Chidiebube, Onyeka, Sunday, & Chiedu, 2025). The primary objectives of this research are: to develop and evaluate machine learning models for predicting maintenance needs in industrial 5-stage compressors; to compare the cost-effectiveness of machine learning-based predictive maintenance models developed; and to identify the key factors that influence the performance of machine learning models in the context of compressor maintenance.

The integration of ML in predictive maintenance not only enhances equipment reliability but also contributes to the broader objectives of Industry 4.0, fostering a more intelligent and responsive industrial ecosystem (Monye, 2023; Okeagu, Nwamekwe, & Nnamani, 2024). Furthermore, this research aims to provide practical recommendations for implementing machine learning-based maintenance strategies in industrial settings, offering industries clear pathways to adopt these advanced technologies effectively. Ultimately, this study seeks to explore the potential of ML in optimizing maintenance strategies, thereby addressing the pressing challenges faced by modern industries (Monye, 2023; Okeagu et al., 2024; Chidiebube, Onyeka, Sunday, & Chiedu, 2025).

This research significantly enhances the existing body of knowledge regarding predictive maintenance in industrial settings, particularly focusing on the application of ML in the maintenance of compressors. Firstly, it addresses a gap in the literature by providing a comprehensive analysis of ML techniques tailored explicitly for these compressors, which have received limited scholarly attention (Igbokwe, Okpala, & Nwamekwe, 2024; Ezeanyim, Ewuzie, Aguh, Nwabueze, & Nwamekwe, 2025). Secondly, the study evaluates the cost-effectiveness of various maintenance strategies, demonstrating how ML can substantially reduce maintenance costs while simultaneously improving equipment reliability (Ringler, 2023; Igbokwe, Nwamekwe, Ono, Nwabunwanne, & Aguh, 2024; Nwamekwe, Ewuzie, Igbokwe, Okpala, & U-Dominic, 2024).

Furthermore, the research proposes a practical framework for implementing predictive maintenance in industrial environments. This framework offers actionable guidelines for practitioners, facilitating the transition from traditional maintenance approaches to more intelligent, data-driven solutions (Okpala, Udu, & Nwamekwe, 2025; Nwamekwe & Nwabunwanne, 2025). By integrating ML into maintenance practices, industries can optimize operational efficiency and minimize unplanned downtime, thereby contributing to the broader objectives of Industry 4.0 (Ezeanyim et al., 2025; Ringler, 2023; Okpala et al., 2025).

1.1. Literature Review

1.1.1. Maintenance Strategies in Industrial Compressors

Maintenance strategies for industrial compressors have undergone significant evolution, transitioning from reactive maintenance, which addresses failures post-occurrence, to preventive maintenance, characterized by scheduled checks and replacements. Reactive maintenance often leads to costly unplanned downtimes, while preventive maintenance can result in unnecessary interventions, thereby escalating operational costs (Townsend & Badar, 2018). In recent years, condition-based maintenance (CBM) has gained traction, where maintenance actions are initiated based on real-time monitoring of specific indicators such as vibration and temperature (Townsend & Badar, 2018; Okpala, Ezeanyim, & Nwamekwe, 2024). This approach aims to enhance reliability and reduce costs by allowing for planned repairs rather than emergency responses (Townsend & Badar, 2018).

However, CBM is not without its limitations. It relies on predefined thresholds that may not adequately capture all factors influencing compressor health, potentially leading to missed fault detections (Okpala et al., 2024). For instance, monitoring systems that detect abnormal behaviours can significantly improve maintenance outcomes, yet they may still overlook critical operational variances (Townsend & Badar, 2018). Furthermore, the integration of advanced technologies, such as machine learning and artificial intelligence, is being explored to enhance fault diagnosis and predictive maintenance capabilities, thereby addressing some of the shortcomings of traditional CBM methods (Nwamekwe & Okpala, 2025; Xiao et al., 2020). Overall, while maintenance strategies for compressors have advanced, ongoing research is essential to optimize these practices further and ensure operational efficiency.

1.1.2. Machine Learning in Predictive Maintenance

Machine learning has emerged as a transformative tool in predictive maintenance, enabling the analysis of extensive datasets to identify patterns indicative of potential failures. By utilizing historical data, machine learning models can forecast when components are likely to fail, facilitating timely maintenance interventions that minimize both downtime and unnecessary maintenance costs (Igbokwe, Okpala, & Nwamekwe, 2024; Okpala, Udu, & Nwamekwe, 2025; Nwamekwe & Okpala, 2025). Various algorithms, including decision trees, support vector machines (SVM), and neural networks, have been effectively applied in this domain, demonstrating promising results in recognizing complex interrelationships among operational variables (Okpala et al., 2025; Nwamekwe & Nwabunwanne, 2025; Raja et al., 2022).

The application of machine learning in predictive maintenance not only enhances the accuracy of failure predictions but also optimizes maintenance schedules, thereby improving operational efficiency (Ezeanyim, Ewuzie, Aguh, Nwabueze, & Nwamekwe, 2025; Ringler, 2023). For instance, studies have shown that hybrid models combining multiple machine learning techniques can outperform traditional methods by leveraging diverse data sources and improving fault detection capabilities (Okpala et al., 2025; Nwamekwe & Igbokwe, 2024). However, challenges remain, particularly in acquiring sufficient labelled data for training supervised models, which is crucial for the reliability of predictions (Igbokwe et al., 2024; Davari et al., 2021). As industries increasingly adopt these advanced methodologies, ongoing research is essential to refine machine learning applications in predictive maintenance, ensuring they meet the dynamic needs of industrial environments (Nwamekwe, Chidiebube, Godfrey, Celestine, & Sunday, 2025).

1.1.3. Previous Studies on Cost-Effective Maintenance

Several studies have investigated the cost-effectiveness of various maintenance strategies, with a notable focus on the advantages of predictive maintenance (PdM). Research indicates that PdM can yield substantial cost savings by minimizing unplanned downtime and optimizing maintenance schedules, thereby enhancing operational efficiency (Wang et al., 2018; Koops, 2018). For instance, Wang et al. (2018) highlight that the cost ratio of scheduled to unscheduled maintenance significantly influences maintenance decisions, suggesting that effective predictive strategies can lead to lower overall costs. Similarly, Koops (2018) emphasizes the potential of accurate failure predictions in reducing unscheduled engine removals, which can have a profound economic impact on operations.

However, the success of predictive maintenance is contingent upon several factors, including the quality of data utilized for model training, the accuracy of predictions, and the timely implementation of maintenance recommendations (Nwamekwe, Chinwuko, & Mgbemena, 2020; Ameeri, 2023). Studies have shown that integrating cost analysis into the development of maintenance strategies is crucial, as it ensures that the chosen approach not only enhances reliability but also provides tangible financial benefits (Alves et al., 2020). For example, predictive maintenance programs have been reported to reduce maintenance costs by 25% to 35% and downtime by 35% to 45%, demonstrating the economic viability of these strategies (Alves et al., 2020). Therefore, a comprehensive understanding of both the technical and financial aspects of predictive maintenance is essential for organizations aiming to optimize their maintenance practices and achieve significant cost reductions (Nwamekwe, Ewuzie, Igbokwe, U-Dominic, & Nwabueze, 2024).

2. Methodology

This experimental study follows the CRISP-DM methodology, as illustrated in Figure 1. CRISP-DM (Cross-Industry Standard Process for Data Mining) is a widely recognized and commonly applied process model for data mining and data science projects. It offers a structured, iterative framework comprising six key phases: Business Understanding, Data Understanding, Data Preparation, Modelling, Evaluation, and Deployment.

This experiment was conducted on Google Colab using Python and libraries like Scikit, Matplotlib, NumPy, Pandas, and Seaborn. Scikit is a versatile library for predictive data analysis and machine learning, offering a variety of tools and algorithms for classification, model evaluation, and more. Matplotlib is used for plotting and visualization, while Pandas is designed for data analysis and manipulation. Seaborn, built on top of Matplotlib, enhances data visualization. The final dataset, Indorama-EPCL, contains 98,794 records.

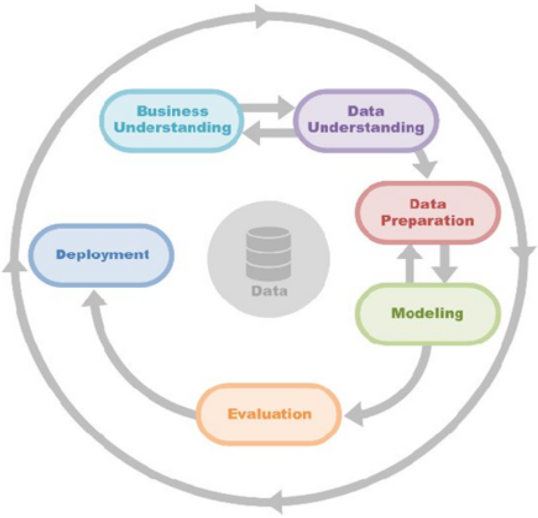

Figure 1: CRISP-DM (data mining framework) [27]

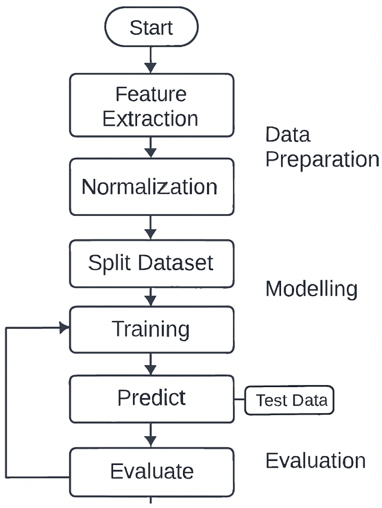

Figure 2: Experimental Flow Diagram

2.1. Data collection

A CSV file containing maintenance datasets was obtained from Indorama-Eleme Petrochemical Ltd. In compliance with a non-disclosure agreement, the variable names within the dataset were recoded to protect the confidentiality of critical equipment attributes in the facility. The datasets include the maintenance history of various rotary equipment in the facility. However, the maintenance history of the gas compressor, K-1, was filtered explicitly from the dataset, as presented in Table 1.

Table 1: First 5 rows of the Maintenance History of the Gas Compressor Dataset 3.2 Exploratory Data Analysis (EDA).

S/N	Date	M2	M3	M4	M5	M6	M7	M8	M9	M10	Failure
1	01,01,2015	2.16E+08	55	0	52	6	407438	0	0	7	0
2	01,02,2015	1650864	56	0	52	6	407438	0	0	7	0
3	01,03,2015	1.24E+08	56	0	52	6	407438	0	0	7	0
4	01,04,2015	1.28E+08	56	0	52	6	407439	0	0	7	0
5	01,05,2015	97393448	56	0	52	6	408114	0	0	7	0

Table 2 presents a detailed summary of the dataset, showing significant variability across the 10 features (M2 to M10) and the target variable "Failure." The consistent total of 98,794 observations reflects a complete dataset, while the wide range of mean values, from 0.000086 to 12.464093, and standard deviations, from 7.032454e+07 to 160.507272, indicate diverse feature distributions.

Additional insights are provided by the minimum, maximum, and percentile values, revealing skewness and the potential presence of outliers in certain features. This comprehensive statistical analysis will inform data preparation, feature engineering, and modelling choices to address the issue at hand effectively.

Table 2: Features Attributes Table.

	Failure	**M2**	**M3**	**M4**	**M5**	**M6**	**M7**	**M8**	**M9**	**M10**
count	98794	98794	98794	98784	98794	98794	98794	98791	98791	98794
mean	0.00086	1.22E+08	166.9919	12.46409	1.891976	13.61847	254441.8	0.290897	0.290897	13.51982
std	0.02932	7.03E+07	2242.806	328.6816	20.73857	14.52715	93644.26	7.921502	7.921502	160.5073
min	0	0	0	0	0	2	0	0	0	0
25%	0	6.15E+07	0	0	0	8	220916.3	0	0	0
50%	0	1.23E+08	0	0	0	10	248626	0	0	0
75%	0	1.83E+08	0	0	0	12	301876	0	0	0
max	1	2.44E+08	64968	80000	1666	98	689161	832	832	10137

The histogram for the feature M10 in Figure 3 shows a skewed distribution, with most values clustered at the lower end and a long tail stretching to the right. The wide range of values, from approximately 0 to 5,000, indicates significant variation in the feature. A noticeable peak suggests a concentration of values within a specific range, likely representing the most common occurrences. The relatively few data points at the higher end of the range suggest that extreme or outlier values for M10 are less common compared to the more frequent, lower-range values.

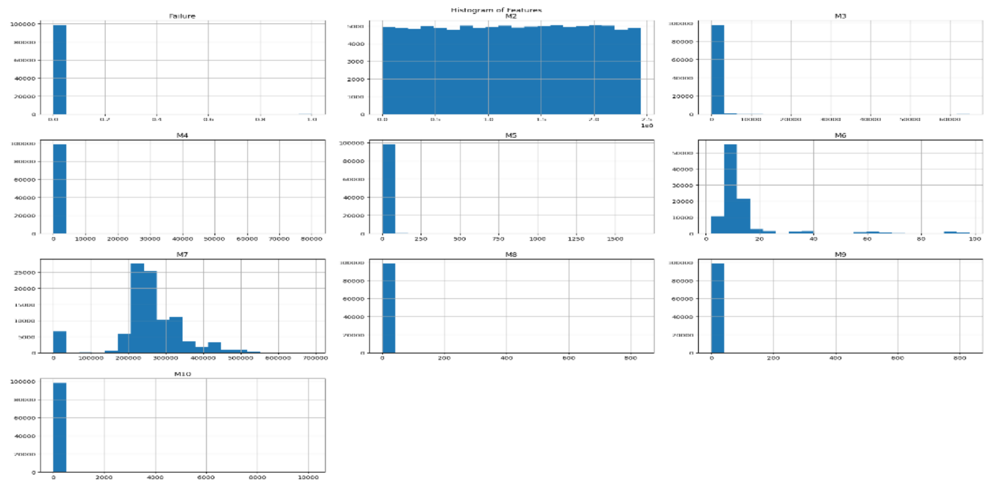

Figure 3. Histogram of features.

Figure 4 provides a clear visual summary of the statistical distribution of the various features in the dataset. The wide range of values and disparate interquartile ranges across the features suggest significant divergency, with some variables exhibiting outliers and skewed distributions. Features like M2 have considerably larger value ranges compared to others. Figure 5 presents a set of scatter plots, each illustrating the statistical distribution and relationships between various features or variables within the dataset. The subplots reveal significant differences in scales, data densities, and patterns. Some exhibit clear linear or curvilinear trends, while others display more scattered, irregular distributions, with potential outliers and anomalies.

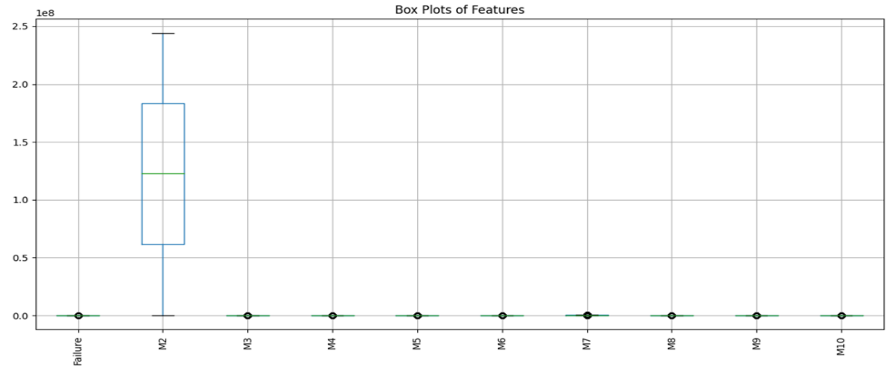

Figure 4. Box plots of features.

Figure 5. Scatter plots of features

Figure 6's correlation matrix offers a succinct statistical summary of the relationships between the features (M1 to M10) and the target variable "Failure" in the dataset. The correlations range from -0.0033 to 1.0, indicating a broad spectrum from weak to strong associations. Key observations include the moderate correlation of M4 and M10 with Failure (0.45). In contrast, features like M1, M2, and M3 show very low correlations, close to 0, suggesting minimal to no linear association with the target variable.

Finally, figure 7, featuring a mix of histograms, scatter plots, time series, and value distributions, provides valuable insights into the diverse characteristics of the features. The histograms and scatter plots highlight the varying ranges, shapes, and bivariate relationships, emphasizing the dataset's heterogeneity. The time series and value plots further illustrate the dynamic nature of the dataset, showing fluctuations, patterns, and skewness over time.

In summary, the histograms display the value distributions across different features, the scatter plots emphasize the intricate bivariate relationships between the features, and the time series plots highlight the data's dynamic characteristics. Together, these visualizations provide essential insights into the dataset's complexity and variability, informing key decisions on preprocessing, feature engineering, and the choice of suitable modelling techniques to accurately capture the underlying patterns and relationships for strong predictive modelling.

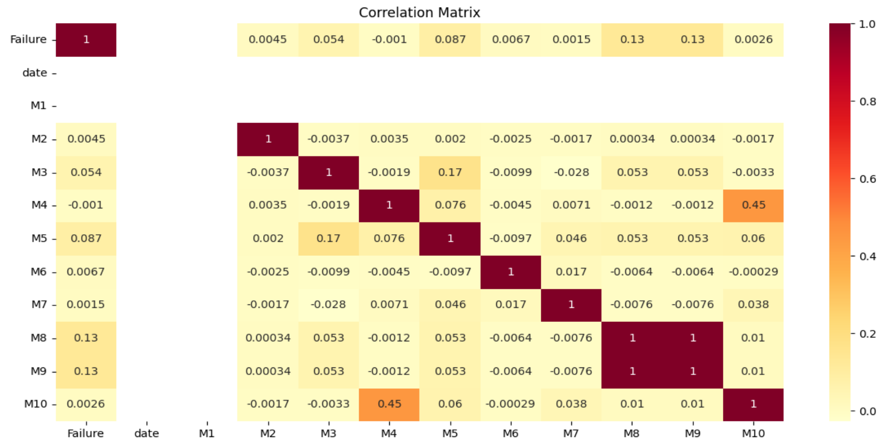

Figure 6. Correlation matrix.

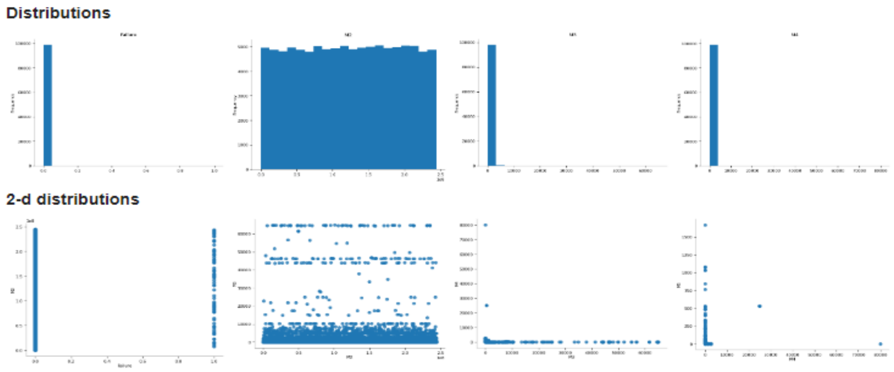

Figure 7. Distributions and time series plot

2.2. Data Preparation

Data preparation typically includes tasks such as handling missing values, addressing data inconsistencies, and restructuring the data to align with the analysis requirements. Addressing potential outliers was a crucial step in the process.

2.3. Data Cleaning

The missing values in the Date column were imputed with the median, as it is a numeric column and the median is a robust central tendency measure, less influenced by outliers than the mean. This makes it an appropriate choice for filling missing values in numeric columns. Furthermore, the numeric features were normalized using the Standard Scaler from scikit-learn, ensuring the features were on a similar scale, which is crucial for many machine learning algorithms. Lastly, the imbalance in the target variable (Failure) was handled using SMOTE (Synthetic Minority Over-sampling Technique) from the imbalanced-learn library. This approach generated synthetic samples for the minority class, helping to balance the target variable, as shown in Figure 8, and improve the performance of the machine learning models.

2.4. Feature Engineering

Feature engineering and dimensionality reduction techniques are commonly used to create new, more insightful features and to reduce the number of variables, respectively. According to Nwamekwe, Ewuzie, Igbokwe, Nwabunwanne, and Ono (2025), feature engineering involves generating new, more informative features from the existing data by applying transformations, combining features, or incorporating domain-specific knowledge, to enhance the predictive accuracy and interpretability of the machine learning model. Dimensionality reduction techniques, on the other hand, focus on minimizing the number of input features while retaining the most critical information, thus improving model training and generalization (Van Der Maaten, Postma, & van den Herik, 2009).

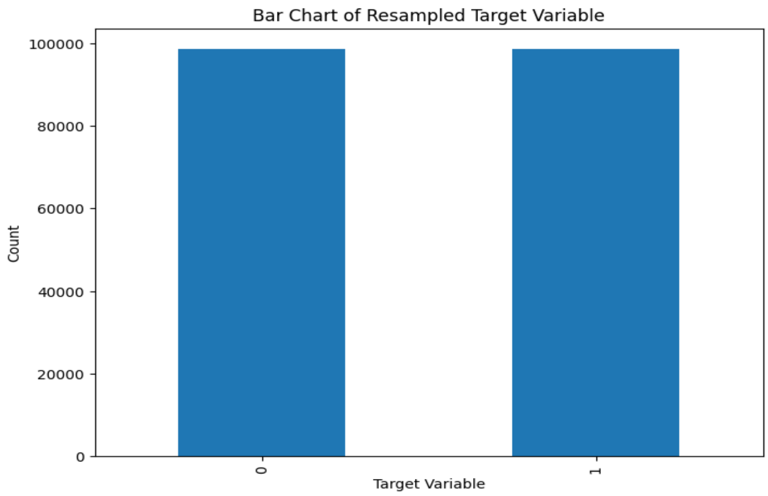

Figure 8: Bar Chart of resamples target variable.

Table 3: Feature ranking statistics

Statistic	Value
Mean	13597.700000
Median	84.500000
Standard Deviation	30178.875675
Minimum	28.000000
Maximum	98322.000000

The feature ranking statistics (Table 3) indicate a dataset with a broad range of feature importance. The average feature count of 13,597.70 suggests that, on average, the features possess a relatively large number of unique values, which could be advantageous for capturing intricate patterns in the data. However, the median feature count of 84.50 is much lower than the mean, indicating a skewed distribution with a small number of features having an exceptionally high count of unique values. This is further evidenced by the significant standard deviation of 30,178.88, reflecting substantial variability in the feature counts. The minimum feature count of 28 and the maximum of 98,322 highlight the wide range of feature importance, from less informative features to highly detailed ones. This variation in feature characteristics suggests that employing careful feature selection or dimensionality reduction methods may be necessary to identify the most relevant and valuable features for the machine learning task at hand.

Table 4: Feature ranking.

No.	Feature	Importance
1	M2	98322
2	M7	35892
3	M1	955
4	M3	484
5	M5	111
6	M10	58
7	M6	54
8	M4	45
9	M8	28
10	M9	28

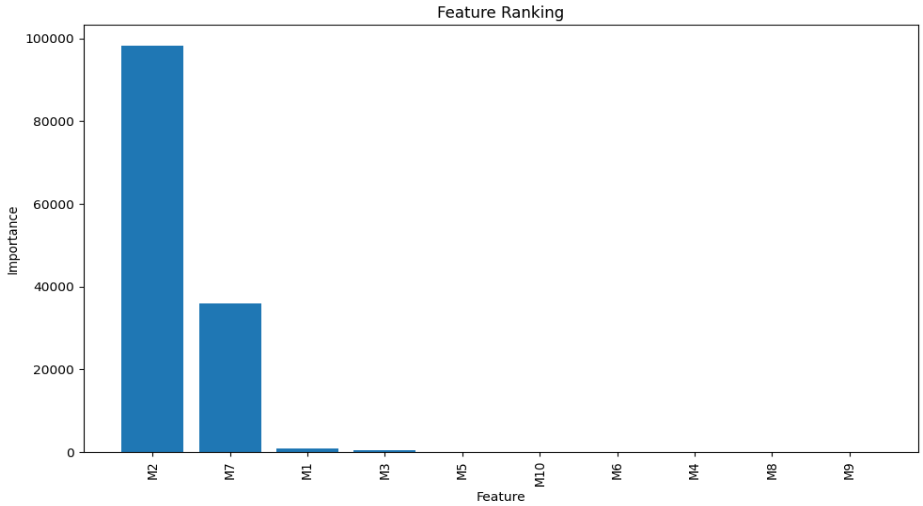

Figure 9. Feature ranking.

The feature ranking in this code, as depicted in Table 4 and Figure 9, is determined by the number of unique values in each feature column of the dataset. The reasoning behind this method is that features with a greater number of unique values are more likely to be informative and significant for the task at hand, as they can capture finer patterns or distinctions in the data. The unique value count for each feature was computed, and the features were then sorted in descending order based on their unique value counts. To identify the best variable, the feature with the highest number of unique values (M2 and M7) was selected, as this likely represents the most informative and discriminative feature for the given maintenance dataset.

2.5. Machine Learning Models

2.5.1. Model Selection

In this study, four machine learning algorithms, namely Random Forest, Logistic Regression, Gradient Boosting, and Decision Trees, were chosen to develop predictive maintenance models for industrial applications. Each of these algorithms brings unique strengths that make them suitable for handling the complexities inherent in industrial maintenance data (Nwamekwe & Nwabunwanne, 2025). By leveraging these algorithms, the research aims to enhance predictive maintenance practices, predict equipment failures, and optimize maintenance schedules to improve operational efficiency and reduce downtime. The study aligns with the growing interest in utilizing machine learning for predictive maintenance in various industries, emphasizing the importance of selecting appropriate algorithms to achieve accurate and reliable maintenance predictions.

Random Forest, an ensemble learning method, is adept at handling large datasets with high-dimensional features by combining multiple decision trees to create a robust model. It effectively mitigates issues like variance and overfitting commonly associated with individual decision trees, making it a valuable tool for predicting maintenance needs in dynamic industrial settings. Leveraging Random Forest in predictive maintenance models can enhance operational efficiency, reduce downtime, and optimize maintenance schedules by accurately predicting equipment failures (Alrabghi & Tiwari, 2015).

Logistic Regression, a widely used statistical model, is valuable for binary classification tasks like predicting maintenance needs due to its simplicity, interpretability, and computational efficiency (Nwamekwe et al., 2025). It is particularly effective in scenarios where model transparency is crucial, providing easily interpretable results (Wang et al., 2018). Leveraging Logistic Regression in predictive maintenance models can enhance decision-making processes, optimize maintenance schedules, and improve operational efficiency by accurately predicting maintenance events (Mobtahej, Zhang, Hamidi, & Zhang, 2022).

Gradient Boosting, an ensemble technique, sequentially builds models by correcting errors from previous models, typically decision trees, to enhance overall performance (Igbokwe, Nwamekwe, Godwin, & Mba, 2025). This method effectively captures complex data relationships, leading to superior predictive accuracy, although it may require careful tuning to prevent overfitting (Igbokwe et al., 2025). By leveraging Gradient Boosting in predictive maintenance models, organizations can improve decision-making processes, optimize maintenance schedules, and enhance operational efficiency by accurately predicting maintenance needs.

Decision Trees are essential models that partition data based on feature values to create an interpretable structure. Although they can be prone to overfitting when used in isolation, their simplicity and interpretability make them valuable in both standalone and ensemble methods. By integrating Decision Trees into predictive maintenance models, organizations can improve decision-making processes, optimize maintenance schedules, and enhance operational efficiency by accurately predicting maintenance needs. This trend of leveraging Decision Trees to boost efficiency and effectiveness in predictive maintenance practices is observed across various industries.

The selection of these models was informed by their ability to address specific predictive maintenance challenges, such as balancing accuracy, interpretability, and computational efficiency. Each model was trained and validated to evaluate its performance, aiming to identify the most effective approach for predicting maintenance requirements in industrial five-stage compressors.

2.6. Model Training

The modelling process involved dividing the dataset into an 80% training set and a 20% test set. The chosen models were trained on a portion of the data, while the rest was used for validation. Cross-validation techniques were applied to assess model performance and prevent overfitting to the training set. Metrics such as accuracy, precision, recall, and the F1 score were used to evaluate the models' effectiveness in predicting maintenance needs, providing a thorough assessment of their predictive abilities. This comprehensive evaluation ensures the predictive maintenance models are reliable and accurate, supporting better decision-making and improved operational efficiency in maintenance.

The selected models were chosen for several reasons. Key benefits include their capacity to handle high-dimensional and noisy datasets effectively, a robust architecture that allows scalability to large data volumes, minimal input preparation, implicit feature selection, strong overall performance, and accessible open-source implementations. Given that the maintenance dataset contains over 90,000 noisy observations, these models were the optimal choice.

During training, the Random Forest model is built iteratively, constructing an ensemble of decision trees. The algorithm continuously adjusts the parameters of each tree to minimize the overall error on the training data. This process continues until the model reaches its optimal fit to the training set. Afterward, the performance of the final Random Forest classifier is tested on the 20% test set to provide an unbiased evaluation of its generalization capabilities, ensuring that the model has not merely memorized the training data. Similar methods were applied to other models as well.

2.7. Evaluation

During the evaluation phase, multiple techniques were used to assess the performance of the Random Forest classifier. This involved cross-validation to ensure the model's performance was tested on various data subsets, along with sampling methods such as train-test splits or k-fold cross-validation. A confusion matrix was then employed to give a detailed analysis of the classifier's predictive abilities, including accuracy, precision, recall, and F1-score.

To fine-tune the model's hyperparameters, the GridSearchCV method from the Scikit-learn library was applied. This approach performs an exhaustive search across a predefined grid of hyperparameter values, automatically training and evaluating the model for each combination. The process is iterative, continuously retraining the model while adjusting hyperparameters until the optimal setup is achieved, maximizing the classifier's performance on the validation data. This structured hyperparameter tuning ensures that the final Random Forest model is thoroughly optimized for the specific dataset and problem at hand.

2.8. Cross Validation

Cross-validation (CV) is a method used to assess how well a model will perform on unseen data. It is a process designed to evaluate the robustness of models. The repeated stratified k-fold cross-validation technique was applied to validate the balanced dataset. This approach is a variation of k-fold cross-validation that uses stratified random sampling to create the folds. The dataset was divided into five folds and repeated three times, with each repeat splitting the training set into one testing fold and four training folds. When applied efficiently, this method can produce accurate results (Souza et al., 2020; Chawla, Bowyer, Hall, & Kegelmeyer, 2002).

2.9. Performance Measurement

A confusion matrix was utilized to assess the performance of the algorithms. As presented in Table 5, the confusion matrix is a table that shows the predicted classifications made by a model in the rows and the actual (true) classifications in the columns. This format provides a clear visualization of the model's performance by showing the number of true positives, true negatives, false positives, and false negatives.

Table 5: Confusion matrix

	Predicted Negative	**Predicted Positive**
Actual Negative	True Negatives	False Positives
Actual Positive	False Negatives	True Positives

True Negatives (TN) refer to the number of correctly identified negative examples, while False Positives (FP) represent incorrectly classified positive examples. False Negatives (FN) denote falsely classified negative examples, and True Positives (TP) indicate the correctly classified positive examples (Chawla, Bowyer, Hall, & Kegelmeyer, 2002).

Accuracy is the proportion of correct predictions out of the total predictions, measured as the ratio of correct classifications to the total classifications. It can be calculated using the equation:

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \qquad \text{(Eqn 1)}$$

Recall is the proportion of correctly predicted positive examples out of all actual positive examples. It is measured as the ratio of actual positive examples to the sum of actual positive and false negative examples. The recall is given by:

$$Recall = \frac{TP}{TP + FN} \qquad \text{(Eqn 2)}$$

Precision represents the proportion of correctly predicted positive examples out of all predicted positive examples. It is measured as the ratio of actual positive examples to the sum of actual positive and false positive examples. Precision is calculated by:

$$Precision = \frac{TP}{TP + FP} \qquad \text{(Eqn 3)}$$

F1-score is the harmonic mean of recall and precision, calculated using the equation:

$$F1 = \frac{2 * Precision * Recall}{Precision + Recall} \qquad \text{(Eqn 4)}$$

Kappa coefficient measures the level of agreement between two raters and is also known as Cohen's kappa, introduced by Jacob Cohen. It can be used for binary and multi-class classifications, with values ranging from -1 (indicating poor prediction) to +1 (perfect prediction) [37]. The kappa is calculated by:

$$Kappa = \frac{2 * (TP * TN - FP * FN)}{(TP + FP) * (FP + TN) + (TP + FN) * (FN + TN)} \qquad \text{(Eqn 5)}$$

The Matthews Correlation Coefficient (MCC), or phi (ϕ) coefficient, is applied to evaluate binary and multi-class classifications. Like kappa, it ranges between -1 and +1, where -1 indicates poor prediction and +1 indicates perfect prediction (Chicco, Warrens, & Jurman, 2021). MCC is calculated as:

$$MCC = \frac{TP * TN - FP * FN}{\sqrt{(TP + FP) * (TP + FN) * (TN + FP) * (TN * FN)}} \qquad \text{(Eqn 6)}$$

The Receiver Operating Characteristic (ROC) curve visualizes the classification performance between the true positive rate (TPR) and the false positive rate (FPR). The area under the ROC curve (ROC AUC) represents the classifier's performance, with a larger area indicating better classification. TPR (also known as Recall or Sensitivity) and FPR (also known as Specificity) are used to plot the ROC curve. They are calculated using the following equations:

$$TPR = \frac{TP}{TP + FN} \qquad \text{(Eqn 7)}$$

$$FPR = \frac{FP}{FP + TN} \qquad \text{(Eqn 8)}$$

2.10. Cost Analysis Framework

This study assessed not only the models' predictive accuracy but also developed a cost analysis framework to evaluate the financial impacts of different maintenance strategies. The framework included direct maintenance costs, unplanned downtime costs, and potential savings from avoiding unnecessary maintenance. By integrating cost analysis with predictive maintenance, the research aimed to provide a complete understanding of the benefits and drawbacks of each strategy.

3. Results and Discussion

3.1. Model Performance Metrics

The performance metrics for the most effective models, Random Forest and Decision Trees, are presented in Table 6 and Figure 10. The Random Forest model also demonstrated remarkable performance across various metrics, boasting an accuracy of 0.9406, a recall of 0.9690, a precision of 0.9169, an F1-score of 0.9422, a Kappa coefficient of 0.8813, and an MCC of 0.8827. These outcomes highlight the Random Forest model's proficiency in accurately classifying instances with high true positive rates and low false positive rates. Furthermore, the Kappa coefficient and MCC values reinforce the model's strong performance by indicating substantial agreement between predicted and actual labels while considering chance agreements. Similarly, the Decision Trees model exhibited robust performance, achieving an accuracy of 0.9298, a recall of 0.9434, a precision of 0.9183, an F1-score of 0.9307, a Kappa coefficient of 0.8595, and an MCC of 0.8599, albeit with slightly lower results compared to the Random Forest model across most metrics.

Table 6: Model Performance Metrics

S/N	Model	Accuracy	Recall	Precision	F1-Score	Kappa	MCC
1	Random Forest	0.9406341809340493	0.9690319310694374	0.9168904661423365	0.9422404021487358	0.8812724195084172	0.8826990368087178
2	Logistic Regression	0.5474116097659811	0.5632539280283831	0.5456643425316704	0.5543196328810853	0.09484060989313337	0.09488986120311084
3	Gradient Boosting	0.7647908013372505	0.8097820577800304	0.7427362744642276	0.7748114740185738	0.5296072231642988	0.5317753399691456
4	Decision Trees	0.9297690203626785	0.9433857070451089	0.9182988800631506	0.9306732668316707	0.859540319814128	0.8598608436048584

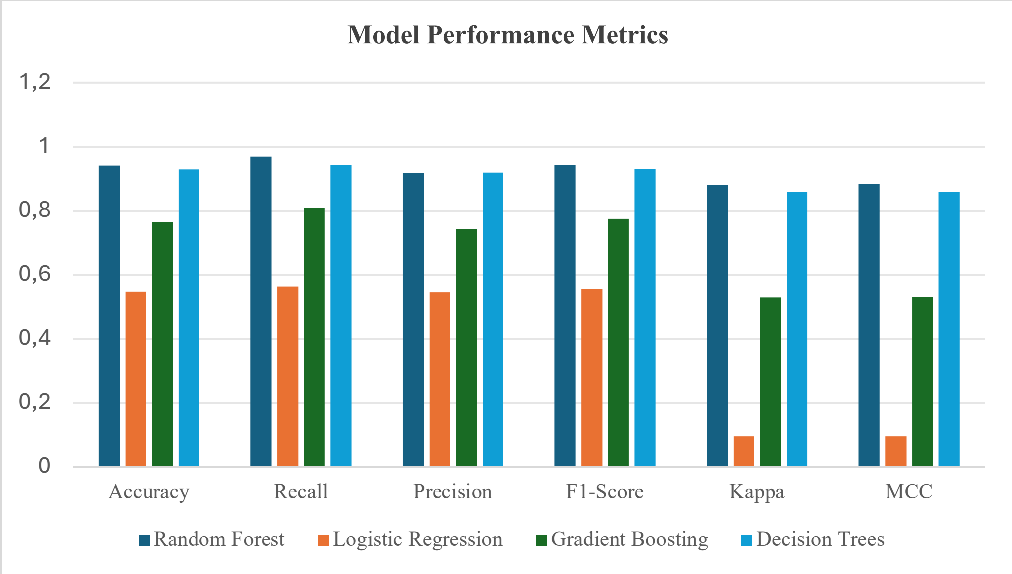

Figure 10. Model Performance Metrics

3.2. Hyperparameter Optimization

The optimal hyperparameters for the Random Forest model were identified as max_depth: 15 and n_estimators: 150, as shown in Table 7. This finding suggests that the model benefited from a deep tree structure and a substantial number of estimators, enabling it to capture intricate data relationships while maintaining strong generalization capabilities.

Table 7. Best Hyperparameters

max_depth	15
n_estimators	150

3.3. Confusion Matrix:

In figure 11, the examination of the confusion matrix allows for a comparative analysis of the classification effectiveness of Random Forest, Logistic Regression, SVM, and Decision Trees. Random Forest shows 19,144 true positives, 18,044 true negatives, 584 false positives, and 1,710 false negatives, indicating high precision and recall. Decision Trees have similar performance with 18,634 true positives, 18,122 true negatives, 1,096 false positives, and 1,632 false negatives. Logistic Regression faces challenges, reporting 11,213 true positives, and 10,511 true negatives, along with significantly elevated false positives (8,517) and false negatives (9,243). SVM holds a middle position with 12,742 true positives, 13,117 true negatives, 6,988 false positives, and 6,637 false negatives. Therefore, Random Forest and Decision Trees demonstrate superior predictive accuracy.

3.4. Cross-Validation Performance

Table 8 and Figure 12 illustrate that the Random Forest model surpassed the other models in terms of cross-validation, exhibiting a mean score of 0.9231528514303886. This outcome signifies the Random Forest model's capability to identify underlying patterns in the data and generalize effectively during cross-validation. Moreover, the Decision Trees model demonstrated respectable performance with a mean cross-validation score of 0.9159954647893519, indicating its suitability for binary classification tasks. Conversely, the Logistic Regression models displayed notably lower cross-validation scores.

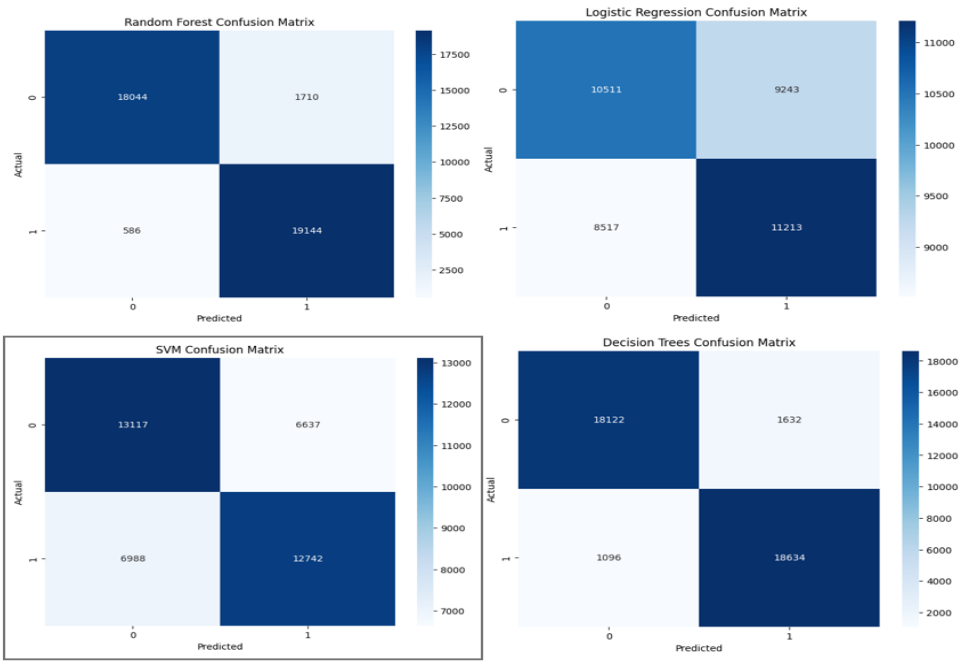

Figure 11. Comparative Confusion Matrix

Table 8. Cross-Validation Performance

		Cross-Validation Scores					
S/N	**Model**	**1**	**2**	**3**	**4**	**5**	**Mean Cross-Validation Scores**
1	Random Forest	0.9276669	0.92080336	0.9313646	0.92227034	0.91365904	0.9231528514303886
2	Logistic Regression	0.53049336	0.53971229	0.53925641	0.53516703	0.47248183	0.5234221844233692
3	Gradient Boosting	0.74181947	0.70854017	0.78267146	0.73717803	0.75151331	0.7443444873434075
4	Decision Trees	0.91880255	0.91414244	0.92237362	0.91419092	0.9104678	0.9159954647893519

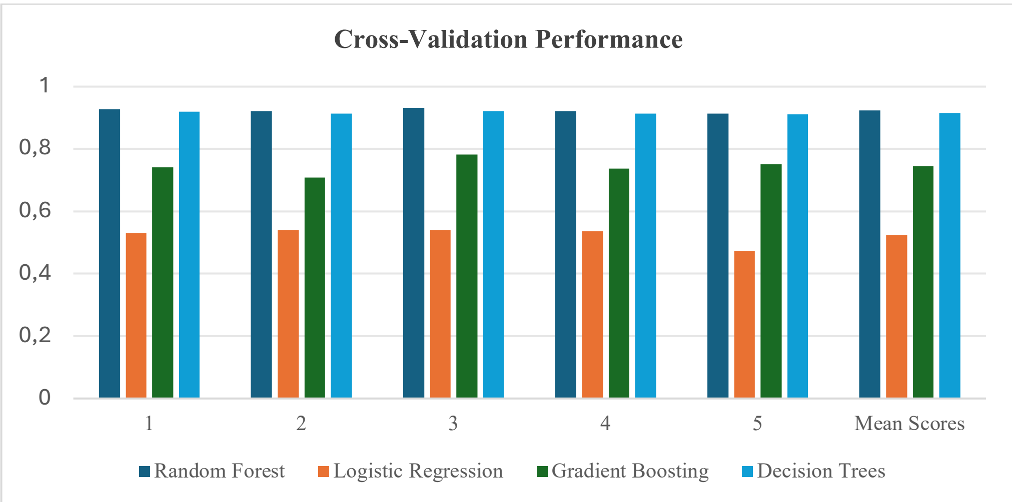

Figure 12. Cross-Validation Performance

3.5. Cost-Effectiveness of Machine Learning Approaches

From Table 9 and Figure 13, the cost analysis showed that implementing machine learning-based predictive maintenance could lead to substantial financial benefits for industrial operators. By optimizing maintenance schedules and reducing downtime, the proposed approach significantly lowered maintenance costs compared to traditional strategies. The study also found that the initial investment in data collection and model development was outweighed by the long-term savings, making predictive maintenance a cost-effective solution for industrial 5-stage compressors.

Among the models that were tested, Random Forest took the longest time to train, clocking in at 46.85 seconds. Gradient Boosting wasn't too far behind, taking 27.16 seconds. In contrast, Decision Trees required significantly less time, just 1.07 seconds, making them a strong contender if you're looking for a balance between speed and accuracy. Logistic Regression, on the other hand, was lightning-fast, with a training time of only 0.088 seconds and a prediction time of 0.00083 seconds. However, its speed comes at a cost: the model didn't perform as well as the others, which means it might not be the best fit for this particular dataset, even though it's very efficient in terms of computational resources.

Table 9. Time Cost Analysis

Model	Training Time (s)	Prediction Time (s)
Random Forest	46.85	1.07
Logistic Regression	0.088	0.00083
Gradient Boosting	27.16	0.058
Decision Trees	1.07	0.0099

4. Discussion

4.1. Comparative Performance of Models

The performance of machine learning models applied in predictive maintenance tasks in this research, reveals significant differences across various metrics, providing valuable insights into model effectiveness and practicality. The Random Forest model consistently outperformed other models, including Decision Trees, Logistic Regression, and Gradient Boosting, across most evaluation metrics. This section will discuss the comparative performance of these models, with a focus on accuracy, recall, precision, F1-score, cross-validation results, time cost, and overall cost-effectiveness in the context of predictive maintenance for industrial 5-stage compressors.

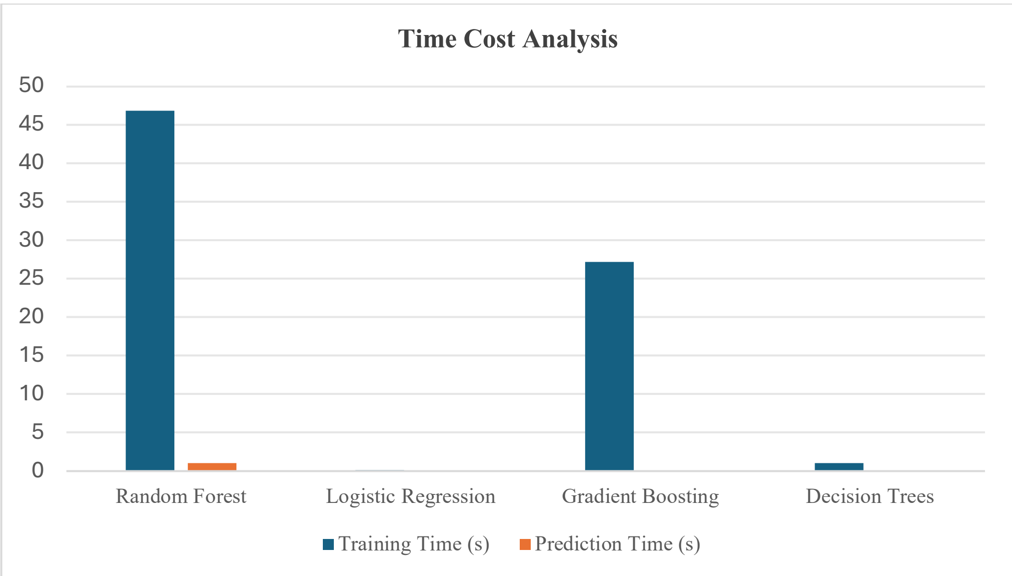

Figure 13. Time Cost Analysis

4.2. Accuracy and Model Classification Performance

From the research, the Random Forest model exhibited the highest accuracy, with a score of 94.06%, surpassing other models such as Decision Trees, Gradient Boosting, and Logistic Regression. The Decision Trees model achieved a strong accuracy of 92.98%, closely following Random Forest but with slightly lower performance across most metrics. In contrast, Gradient Boosting and Logistic Regression demonstrated significantly lower accuracy, at 76.47% and 54.74%, respectively.

The superiority of Random Forest is further supported by its recall of 96.90% and precision of 91.69%, which indicates its high actual positive rate and ability to minimize false positives. These results suggest that Random Forest is highly effective in identifying potential equipment failures while maintaining reliability. Decision Trees, while similarly robust with a recall of 94.34% and precision of 91.83%, still fall short of Random Forest's performance. Gradient Boosting, with a recall of 80.97% and precision of 74.27%, and Logistic Regression, with a recall of 56.32% and precision of 54.57%, were significantly less effective, making them less suitable for high-stakes predictive maintenance tasks.

4.3. F1-Score, Kappa, and MCC

F1-scores, Kappa coefficients, and MCC are vital indicators of model performance in imbalanced datasets, where precision and recall need to be balanced. The Random Forest model's F1-score of 0.9422, Kappa coefficient of 0.8813, and MCC of 0.8827 highlight its exceptional balance between precision and recall, ensuring reliability in real-world predictive maintenance scenarios. Decision Trees performed slightly lower, with an F1-score of 0.9307, a Kappa coefficient of 0.8595, and an MCC of 0.8599, which still demonstrate strong agreement between predicted and actual labels but fall short of Random Forest's levels.

On the other hand, Gradient Boosting (F1-score of 0.7748, Kappa coefficient of 0.5296, MCC of 0.5317) and Logistic Regression (F1-score of 0.5543, Kappa coefficient of 0.0948, MCC of 0.0949) exhibited weaker results, making them less desirable for predictive maintenance when high reliability is required. The comparatively lower F1-score for these models is due to their struggle with accurately capturing both false positives and false negatives in the dataset.

4.4. Cross-Validation Results

Cross-validation scores provide insight into how well the models generalize to unseen data. From the above results, Random Forest demonstrated the highest mean cross-validation score of 92.32%, further confirming its capacity to maintain strong generalization performance across different data splits. Decision Trees followed closely, with a mean cross-validation score of 91.60%, further affirming its potential as an alternative model when computational efficiency is prioritized.

Gradient Boosting’s cross-validation score of 74.43% and Logistic Regression's 52.34% significantly lagged, emphasizing their lack of generalization in predictive maintenance tasks. These results reinforce the notion that Random Forest and Decision Trees are better suited for handling complex industrial datasets in the context of maintenance predictions.

4.5. Time Cost Analysis

While Random Forest emerged as the most accurate model, it came at the cost of longer training times. From the results above, Random Forest required 46.85 seconds for training, while Gradient Boosting took 27.16 seconds. Decision Trees were significantly faster, training in just 1.07 seconds, and Logistic Regression was the quickest, with a training time of only 0.088 seconds. This disparity in training time indicates a trade-off between model complexity and speed.

Though Decision Trees provided slightly lower performance than Random Forest, their reduced training time makes them a practical choice for real-time predictive maintenance applications, where faster model updates may be necessary. Logistic Regression, while the quickest model, demonstrated poor predictive accuracy, making it unsuitable for applications where prediction quality is critical, despite its computational efficiency.

4.6. Cost-Effectiveness of Predictive Maintenance

From a cost-effectiveness perspective, as analysed in this research, machine learning-based predictive maintenance showed substantial financial benefits, with Random Forest being the most effective in reducing unplanned downtime and optimizing maintenance schedules. While the initial investment in data collection and model development was significant, the long-term cost savings ranging from 25% to 35% for maintenance costs and 35% to 45% for downtime outweighed these upfront costs. Decision Trees, due to their balance between speed and accuracy, also emerged as a cost-effective solution, especially when computational resources are constrained.

In contrast, Logistic Regression's poor predictive performance and Gradient Boosting's relatively slower training times make them less appealing from a cost perspective, despite their faster prediction times. The findings from the research further highlight that the Random Forest and Decision Trees models are not only technically superior but also financially viable in predictive maintenance applications.

In summary, Random Forest consistently demonstrated superior performance across key metrics, including accuracy, precision, recall, F1-score, and generalization capabilities. Decision Trees followed closely, offering a faster, more computationally efficient alternative, albeit with slightly lower performance. Gradient Boosting and Logistic Regression, while faster in training and prediction, performed significantly worse in terms of classification accuracy and generalization, making them less suitable for predictive maintenance in industrial applications. The comparative analysis in this research emphasizes the importance of selecting models like Random Forest and Decision Trees for high-stakes, cost-effective predictive maintenance strategies in industrial settings.

5. Conclusion

The study applied machine learning (ML) techniques, Random Forest, Decision Trees, Gradient Boosting, and Logistic Regression, to predictive maintenance for industrial 5-stage compressors, comparing model performance, cost-effectiveness, and real-world implications. Random Forest achieved the highest accuracy (94%), outperforming others, while Decision Trees performed slightly lower but still robust; Logistic Regression was the least accurate despite its efficiency. Hyperparameter optimization improved model accuracy, and cost analysis showed significant downtime reduction and maintenance cost savings, making ML-based predictive maintenance cost-effective and scalable. The approach aligns with Industry 4.0 goals, enhancing reliability and operational efficiency. Limitations included reliance on high-quality labelled data, a limited range of models tested, and significant computational costs for complex algorithms like Random Forest. The focus on 5-stage compressors may limit applicability to other equipment. Future research should explore advanced algorithms (deep learning, reinforcement learning), improve data quality through cleaning and augmentation, leverage semi-supervised learning for unlabelled data, and broaden application to diverse industrial contexts for broader generalizability.

Reference

Monye, S. (2023). Overview and impact of maintenance process in 4th industrial revolution. E3s Web of Conferences, 430, 01220. https://doi.org/10.1051/e3sconf/202343001220

Okeagu F. N., Charles Onyeka Nwamekwe, Blessing Precy Nnamani. (2024). Challenges and Solutions of Industrial Development in Anambra State, Nigeria. Iconic Research and Engineering Journals, 7(11), 467-472. https://www.irejournals.com/formatedpaper/1705825.pdf

Chidiebube, I. N., Onyeka, N. C., Sunday, A. P., & Chiedu, E. O. (2025). A Comparative Analysis of Machine Learning Models for Inventory Demand Forecasting in A Food Manufacturing Sme. Indonesian Journal of Innovation Science and Knowledge, 2(3), 35-48. Retrieved from https://knowledge.web.id/index.php/ijisk/article/view/177

Igbokwe, N. C., Okpala, C. C., & Nwamekwe, C. O. (2024). The Implementation of Internet of Things in the Manufacturing Industry: An Appraisal. International Journal of Engineering Research and Development, 20(7), 510-516. https://www.ijerd.com/paper/vol20issue7/2007510516.pdf

Ezeanyim, O. C., Ewuzie, N. V., Aguh, P. S., Nwabueze, C. V., and Nwamekwe, C. O. (2025). Effective Maintenance of Industrial 5-Stage Compressor: A Machine Learning Approach. Gazi University Journal of Science Part A: Engineering and Innovation, 12(1), 96-118. https://dergipark.org.tr/en/pub/gujsa/issue/90827/1646993

Ringler, N. (2023). Machine learning based real time predictive maintenance at the edge for manufacturing systems: a practical example. https://doi.org/10.1109/globconet56651.2023.10150033

Igbokwe, N. C., Nwamekwe, C. O., Ono, C. G., Nwabunwanne, E. C., & Aguh, P. S. A. (2024). The Role of Digital Twins in Optimizing Renewable Energy Utilization and Energy Efficiency in Manufacturing. Siber International Journal of Digital Business (SIJDB), 1(4), 93–111. https://doi.org/10.38035/sijdb.v1i4.262

Nwamekwe, C. O., Ewuzie, N. V., Igbokwe, N. C., Okpala, C. C., and U-Dominic, C. M. (2024). Sustainable Manufacturing Practices in Nigeria: Optimization and Implementation Appraisal. Journal of Research in Engineering and Applied Sciences, 9(3). https://qtanalytics.in/journals/index.php/JREAS/article/view/3967

Okpala, C. C., Udu, C. E., & Nwamekwe, C. O. (2025). Artificial Intelligence-Driven Total Productive Maintenance: The Future of Maintenance in Smart Factories. International Journal of Engineering Research and Development (IJERD), (21)1, 68-74. https://www.ijerd.com/paper/vol21-issue1/21016874.pdf

Nwamekwe C. O., and Nwabunwanne E. C. (2025). Circular Economy and Zero-Energy Factories: A Synergistic Approach to Sustainable Manufacturing. Journal of Research in Engineering and Applied Sciences (JREAS), 10(1), 829-835. https://qtanalytics.in/journals/index.php/JREAS/article/view/4567

Townsend, J. and Badar, M. (2018). Impact of condition monitoring on reciprocating compressor efficiency. Journal of Quality in Maintenance Engineering, 24(4), 529-543. https://doi.org/10.1108/jqme-06-2017-0040

Okpala, C. C., Ezeanyim, O. C., & Nwamekwe, C. O. (2024). The Implementation of Kaizen Principles in Manufacturing Processes: A Pathway to Continuous Improvement. International Journal of Engineering Inventions, 13(7), 116-124. https://www.ijeijournal.com/papers/Vol13-Issue7/1307116124.pdf

Nwamekwe, C. O., and Okpala, C. C. (2025). Circular economy strategies in industrial engineering: From theory to practice. International Journal of Multidisciplinary Research and Growth Evaluation, 6(1): 1773-1782. https://www.allmultidisciplinaryjournal.com/uploads/archives/20250212103754_MGE-2025-1-288.1.pdf

Xiao, S., Ang, N., Zhang, Z., Liu, S., Song, M., and Zhang, H. (2020). Fault diagnosis of a reciprocating compressor air valve based on deep learning. Applied Sciences, 10(18), 6596. https://doi.org/10.3390/app10186596

Okpala C. C., Chukwudi Emeka Udu, & Charles Onyeka Nwamekwe. (2025). Sustainable HVAC Project Management: Strategies for Green Building Certification. International Journal of Industrial and Production Engineering, 3(2), 14-28. https://journals.unizik.edu.ng/ijipe/article/view/5595

Nwamekwe, C. O., and Okpala, C. C. (2025). Machine learning-augmented digital twin systems for predictive maintenance in highspeed rail networks. International Journal of Multidisciplinary Research and Growth Evaluation, 6(1), 1783-1795. https://www.allmultidisciplinaryjournal.com/uploads/archives/20250212104201_MGE-2025-1-306.1.pdf

Raja, H., Kudelina, K., Asad, B., Vaimann, T., Kallaste, A., Rassõlkin, A., ... and Khang, H. (2022). Signal spectrum-based machine learning approach for fault prediction and maintenance of electrical machines. Energies, 15(24), 9507. https://doi.org/10.3390/en15249507

Nwamekwe, C. O., and Igbokwe, N. C. (2024). Supply Chain Risk Management: Leveraging AI for Risk Identification, Mitigation, and Resilience Planning. International Journal of Industrial Engineering, Technology & Operations Management, 2(2), 41–51. https://doi.org/10.62157/ijietom.v2i2.38

Davari, N., Veloso, B., Costa, G., Pereira, P., Ribeiro, R., and Gama, J. (2021). A survey on data-driven predictive maintenance for the railway industry. Sensors, 21(17), 5739. https://doi.org/10.3390/s21175739

Nwamekwe, C. O., Chidiebube, I. N., Godfrey, O. C., Celestine, N. E., & Sunday, A. P. (2025). Resilience and Risk Management in Social Robot Systems: An Industrial Engineering Perspective. Culture Education and technology research (Cetera), 2(2), 1–12. Retrieved from https://cetera.web.id/index.php/ctr/article/view/154

Wang, Y., Gogu, C., Binaud, N., Bès, C., Haftka, R., and Kim, N. (2018). Predictive airframe maintenance strategies using model-based prognostics. Proceedings of the Institution of Mechanical Engineers Part O Journal of Risk and Reliability, 232(6), 690-709. https://doi.org/10.1177/1748006x18757084

Koops, L. (2018). Roc-based business case analysis for predictive maintenance – applications in aircraft engine monitoring. PHM Society European Conference, 4(1). https://doi.org/10.36001/phme.2018.v4i1.460

Nwamekwe, C. O., Chinwuko, C. E. & Mgbemena, C. E. (2020). Development and Implementation of a Computerised Production Planning and Control System. UNIZIK Journal of Engineering and Applied Sciences, 17(1), 168-187

Ameeri, T. (2023). Analysing effective and ineffective impacts of maintenance strategies on electric power plants: a comprehensive approach. Energies, 16(17), 6243. https://doi.org/10.3390/en16176243

Alves, F., Badikyan, H., Moreira, H., Azevedo, J., Moreira, P., Romero, L., ... and Leitão, P. (2020). Deployment of a smart and predictive maintenance system in an industrial case study. https://doi.org/10.1109/isie45063.2020.9152441

Nwamekwe, C. O., Ewuzie, N. V., Igbokwe, N. C., U-Dominic, C. M., & Nwabueze, C. V. (2024). Adoption of Smart Factories in Nigeria: Problems, Obstacles, Remedies and Opportunities. International Journal of Industrial and Production Engineering, 2(2). Retrieved from https://journals.unizik.edu.ng/ijipe/article/view/4167

Tounsi, Youssef and Anoun, Houda & Hassouni, Larbi. (2020). CSMAS: Improving Multi-Agent Credit Scoring System by Integrating Big Data and the new generation of Gradient Boosting Algorithms. 1-7. 10.1145/3386723.3387851.

Nwamekwe, C. O., Ewuzie, N.V., Igbokwe, N. C., Nwabunwanne, E. C., & Ono, C. G. (2025). Digital Twin-Driven Lean Manufacturing: Optimizing Value Stream Flow. Letters in Information Technology Education (LITE), 8 (1), pp.1-13. https://hal.science/hal05127340/

Van Der Maaten, L., Postma, E. O., and van den Herik, H. J. (2009). Dimensionality reduction: A comparative review. Journal of Machine Learning Research, 10(66-71), 13.

Alrabghi, A. and Tiwari, A. (2015). State of the art in simulation-based optimisation for maintenance systems. Computers & Industrial Engineering, 82, 167-182. https://doi.org/10.1016/j.cie.2014.12.022

Nwamekwe, C. O., Ewuzie, N. V., Okpala, C. C., Ezeanyim, C., Nwabueze, C. V., & Nwabunwanne, E. C. (2025). Optimizing Machine Learning Models for Soil Fertility Analysis: Insights from Feature Engineering and Data Localization. Gazi University Journal of Science Part A: Engineering and Innovation, 12(1), 36-60. https://dergipark.org.tr/en/pub/gujsa/issue/90827/1605587

Mobtahej, P., Zhang, X., Hamidi, M., and Zhang, J. (2022). An lstm-autoencoder architecture for anomaly detection applied on compressors audio data. Computational and Mathematical Methods, 2022, 1-22. https://doi.org/10.1155/2022/3622426

Igbokwe, N. C., Nwamekwe, C. O., Godwin, H. C., & Mba W. (2025). Optimization of Overall Equipment Effectiveness Factors in A Food Manufacturing Small and Medium Enterprise. Journal of Research in Engineering and Applied Sciences (JREAS), 10(1), 836-845. https://qtanalytics.in/journals/index.php/JREAS/article/view/4660

Souza, D. V., Santos, J. X., Vieira, H. C., Naide, T. L., Nisgoski, S., and Oliveira, L. E. S. (2020). An automatic recognition system of Brazilian flora species based on textural features of macroscopic images of wood. Wood Science and Technology. Retrieved from https://doi.org/10.1007/s00226-020-01196-z

Nwamekwe, C. O., Okpala, C. C., and Okpala, S. C., (2024). Machine Learning-Based Prediction Algorithms for the Mitigation of Maternal and Fetal Mortality in the Nigerian Tertiary Hospitals. International Journal of Engineering Inventions, 13(7), PP: 132-138. https://www.ijeijournal.com/papers/Vol13-Issue7/1307132138.pdf

Chawla, N. V., Bowyer, K. W., Hall, L. O., and Kegelmeyer, W. P. (2002). SMOTE: Synthetic Minority Over-sampling Technique. Journal of Artificial Intelligence Research. Retrieved from https://doi.org/10.48550/arXiv.1106.1813

Chicco, D., Warrens, M. J., and Jurman, G. (2021). The Matthews Correlation Coefficient (MCC) is More Informative Than Cohen's Kappa and Brier Score in Binary Classification Assessment. IEEE. Retrieved from https://doi.org/10.1109/ACCESS.2021.3084050.